Trends in Biotechnology

CellPress REVIEWS

Review

Synthetic Biological Circuits within an Orthogonal Central Dogma

Alan Costello [1] and Ahmed H. Badran [1,*]

Synthetic biology strives to reliably control cellular behavior, typically in the form of user-designed interactions of biological components to produce a predetermined output. Engineered circuit components are frequently derived from natural sources and are therefore often hampered by inadvertent interactions with host machinery, most notably within the host central dogma. Reliable and predictable gene circuits require the targeted reduction or elimination of these undesirable interactions to mitigate negative consequences on host fitness and develop context-independent bioactivities. Here, we review recent advances in biological orthogonalization, namely the insulation of researcher-dictated bioactivities from host processes, with a focus on systematic developments that may culminate in the creation of an orthogonal central dogma and novel cellular functions.

Highlights

Development of fully synthetic nucleobase pairs that faithfully interact in living cells, and their applications in creating semisynthetic organisms with expanded and orthogonal information-carrying capacity.

Harnessing naturally occurring and mutually orthogonal DNA replication systems to enable replication of target genes. Highly error-prone variations on these systems enable robust directed evolution of biomolecules.

Engineering and directed evolution of mutually orthogonal transcription factors that operate with high dynamic range, low background, and respond to a wide repertoire of stimuli *in vivo*.

Recent developments in *in vivo* orthogonal protein translation including: orthogonal RBS–orthogonal anti-RBS pairs, covalently linked rRNA subunits to discover novel enzymatic capabilities, improved incorporation of non-canonical amino acids and decoding quadruplet codons.

Engineering Gene Circuits

Natural gene circuits (see Glossary) rely on regulatory factors to govern cellular activities, leveraging elements of both circuit insulation and interaction to afford the necessary responses [1]. These natural gene circuits are maintained through discrete functional modules, which allow for both robustness and plasticity in adapting to changing environmental conditions [2]. Over the past two decades, components of natural gene circuits have been isolated and repurposed to develop increasingly complex **engineered gene circuits** [3], wherein multiple user-defined inputs may act in concert to produce the desired outputs [4]. Increasing circuit complexity has been driven by the incremental realization of synthetic biology goals: controlled gene expression [5], **genetic code expansion** [6] and development of **synthetic cells** [7]. In turn, these and other accomplishments have catalyzed significant research into more complex genetic circuitry that continue to co-opt **host** cell machinery and resources for researcher-dictated bioactivities.

Historically, engineered gene circuits have been implemented at the transcriptional and post-transcriptional levels [4,5,8–11], integrating RNAs [9–11], proteins [5,8], or combinations thereof [12]. Engineered circuits, therefore, rely heavily on elements within the central dogma of the host to generate the necessary biomolecules, which can inadvertently reduce host fitness by depleting essential resources or by enforcing non-native functions onto pre-existing machinery [12–19]. Efforts to improve engineered circuit performance and overcome pleiotropic consequences on the host have focused on limiting component overproduction by affecting transcription or translation rates [12]. While this is an important starting point for circuit optimization, these *ad hoc* approaches may similarly disrupt the balance of native cellular components [14]. Furthermore, empirical approaches for improving synthetic circuits and their components become particularly challenging when circuits grow in complexity, as the combinations of possible modifications grows exponentially [13].

To date, bottlenecks in engineering more complex gene circuits have stemmed from two broad factors: (i) engineered circuits typically repurpose components that are optimized to function

[1]Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA

*Correspondence: ahbadran@broadinstitute.org (A.H. Badran).

within their native contexts, and which are not immediately suited for user-defined objectives, such as high levels of basal expression in an off-state [20]; and (ii) their performance is often limited by incompatibility of components, such as crosstalk or toxicity in the chosen host [21]. Insulation of these components from host processes (**biological orthogonalization**; Box 1) has been described as early as the 1960s [22] (Figure 1, Key Figure) and has significantly improved the predictability of engineered circuits by limiting known instances of cellular crosstalk [5]. An orthogonal and user-controlled paralog [23] of the central dogma may therefore continue to improve the reliability of engineered circuits, and enable increasingly numerous elements to coexist alongside host components without adverse cellular effects. In this review, we discuss prior studies aimed at segregating synthetic and host elements, and take inspiration from native biological systems to selectively partition, compartmentalize, or orthogonalize native and/or engineered components to enable specific goals. As modern engineered gene circuits continue to rely heavily on the host central dogma, we discuss these developments with a focus on their integration to furnish an orthogonal central dogma, wherein elements of information maintenance, transfer, and translation have been modified and/or insulated from host interactions.

Orthogonal Genetic Information Storage and Replication

Insulation via Non-canonical Nucleobases

Engineered gene circuits can be introduced into a host chassis as stand-alone episomes [24] or integrated into the host genome [25]. As inadvertent recognition of engineered sequences by host components may affect circuit performance in unpredictable ways, mechanisms that insulate exogenous genetic information from host machinery would limit reliance on the host for replication or gene expression [26]. Epigenetic modification of nucleic acids is among the most common mechanisms exploited by natural biological systems to insulate specific cellular processes. Taking inspiration from these natural systems, Khalil and colleagues recently established an orthogonal methodology to control gene transcription in eukaryotic cells through epigenetic signaling [27]. This system uses a modified nucleobase commonly found in prokaryotes but absent in eukaryotic genomes, N6-methyldeoxyadenosine (m6dA), by porting the requisite methyltransferase and transcription factors to enable efficient and orthogonal information storage and propagation (Figure 2A).

A diversity of DNA modifications exists in natural genomes [28] (Figure 2B), wherein the abundance of non-canonical nucleobases may make their genetic information innately orthogonal to the host machinery. For example, the genomes of a eukaryotic multicellular organism may be identical in sequence but produce distinct gene expression patterns through alternative epigenetic profiles [27]. Such modifications may also provide protective functions to cellular circuits. Bacteriophage genomes can have partial or complete substitution of nucleobases by modified variants [26], which are thought to resist endonuclease-mediated digestion upon infection.

Box 1. Biological Orthogonalization

The term orthogonal or orthogonality in synthetic biology describes the inability of two or more biomolecules, similar in composition and/or function, to interact with one another or affect their respective substrates. For example, two proteases may be mutually orthogonal if they are unable to cleave one another's respective substrates, or two aaRSs that do not cross aminoacylate their noncognate tRNA. By necessity, all biomolecules within a cell should not interact directly with one another. This allows key processes such as those within the central dogma to perform discrete functions (replication, transcription, and translation) while remaining intrinsically linked as each function is dependent on the prior stages. As it relates to the development of engineered gene circuits, the necessary degree of orthogonality depends on the user-defined objectives. For example, it may be necessary to orthogonally control the transcription of multiple genes using independent promoters in an engineered gene circuit. Here, TFs with known mutual orthogonality in their respective operator sequences may be sufficient. Conversely, more complex goals such as the large-scale implementation of expanded genetic codes will undoubtedly require a larger repertoire of orthogonal elements.

Glossary

Aminoacyl-tRNA synthase: also known as a tRNA ligase, is an enzyme that covalently attaches an amino acid to a tRNA.
Biological orthogonalization: purposeful insulation of biomolecules from host components, aimed at limiting undesirable crosstalk with cellular biomolecules or environmental stimuli.
Codon: sequence of three nucleotides in a DNA or RNA molecule that forms a translatable unit within the genetic code.
DNA polymerase: enzyme that polymerizes deoxyribonucleotides to afford DNA.
Elongation factor P: a prokaryotic factor that improves the efficiency of initial peptide bond formation during the transition from initiation to elongation.
Elongation factor Tu: a prokaryotic factor that regulates aaRS selection *en route* to the ribosome.
Engineered gene circuit: synthetic collective of interactions using naturally derived, engineered and/or evolved biomolecules (DNA, RNA, proteins, and/or small molecules) that cumulatively affect a user-defined process or outcome.
Genetic code expansion: addition of synthetic nucleic acids or ncAAs to living biological systems to supplement their information-carrying capacities, often via an increased number of unique and functional codons.
Host: as it pertains to synthetic biology, a host is the cell in which exogenous genetic information has been introduced.
mRNA: contains an open reading frame encoding a protein.
Natural gene circuit: natural collective of interactions using cellular factors (DNA, RNA, proteins, and/or small molecules) that exist within natural biological systems and are integrated to affect a specific process or outcome.
Non-canonical amino acid: unnatural amino acid not found in the 20 proteogenic amino acids commonly found in biological systems.
Orthogonal translation: partitioning of ribosomal pools within a cell, wherein a discrete population of the ribosomes exclusively initiate the translation of researcher-defined orthogonal mRNAs.
Quadruplet codon: sequence of four nucleotides in a DNA or RNA molecule that forms a translatable unit within an expanded genetic code.

Key Figure

Timeline of Developing Orthogonal Elements

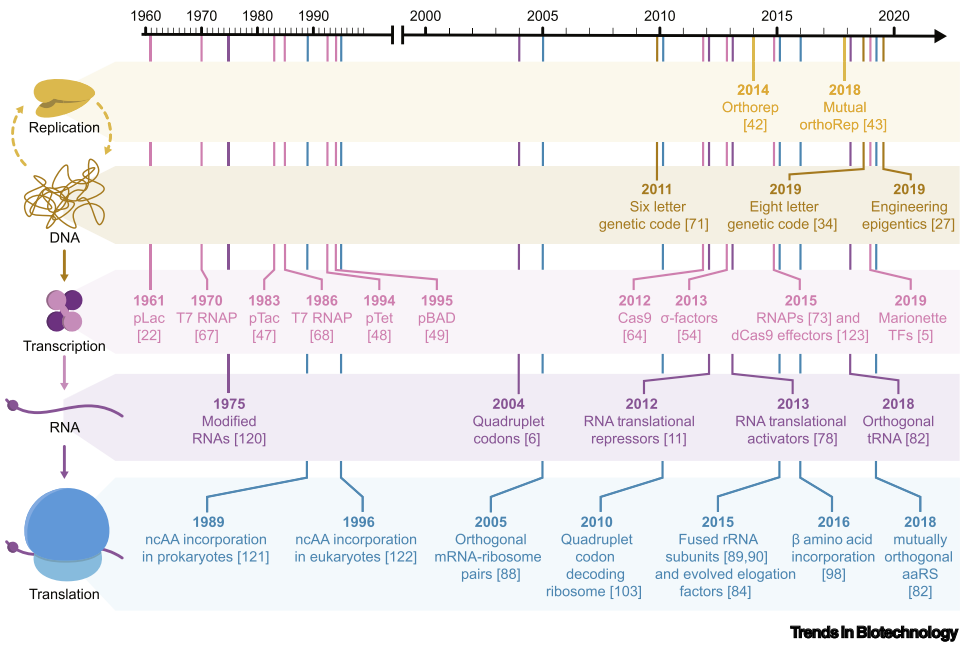

Figure 1. Schematic representation of key orthogonal elements discovered through investigation of natural biological systems, biomolecular engineering, and/or directed evolution. Elements are segregated by color according to their order within the central dogma: genetic information storage and replication (yellow/orange), transcription (pink/purple) and translation (blue) [120–122]. In each case, the year of their first literature report is provided. Abbreviations: aaRS, aminoacyl-tRNA synthetase; ncAA, non-canonical amino acid.

RNA polymerase: enzyme that polymerizes ribonucleotides to afford RNA during transcription, often uses a DNA sequence as the template.
Synthetic cell: broad term that includes an artificial/engineered particle that imitates one or many functions of a natural biological cell, typically defined as biologically active material enclosed in a membrane.
Synthetic nucleobases: synthesized *in vitro* and not found in natural biological systems.
T7: bacteriophage T7 is a lytic virus that infects bacteria.
tRNA: noncoding 70–90 nucleotide RNA molecule that decodes codons in mRNAs and is used by the ribosome to translate mRNA sequences to the corresponding proteins.
Untranslated region: sequence of mRNA flanking the open reading frame defining the protein sequence.

However, incorporation of non-canonical nucleobases may require the action of dedicated polymerases to facilitate replication and propagation [29]. Modified nucleotides in bacteriophages are generated primarily through enzymatic modification of cellular precursors [26]; however, some bacteriophage components can introduce secondary nucleotide modifications following replication [26,30]. It may therefore be possible to exploit these modifications to program additional layers of orthogonal information maintenance in non-natural hosts.

The observation of non-canonical nucleobases in natural systems has motivated significant interest in incorporating synthetic nucleotide pairs into cellular DNA, where such efforts could enable genetic code expansion through increased sequence diversity. Recent work to expand the genetic code has recently resulted in accretion from the canonical four to six [31–33] or eight [34] basepairs using **synthetic nucleobases**, thereby increasing the possible information density while reducing the potential for nucleobase–host component interactions (Figure 2B). While these systems have expanded the repertoire of viable **codons** *in vivo* for **non-canonical amino acid** (ncAA) incorporation [33], they require *in vitro* synthesis of (deoxy)nucleoside triphosphates for cellular uptake to maintain non-canonical information transfer [35]. Similarly, non-canonical sugars in nucleic acids have served as a basis for artificial information maintenance and transfer using evolved polymerases [36]. Finally, nucleic acid phosphate backbones are known to be extensively modified to phosphorothioate across diverse bacterial genera [37]; an interesting example of an artificial innovation that was later discovered to occur naturally in biological

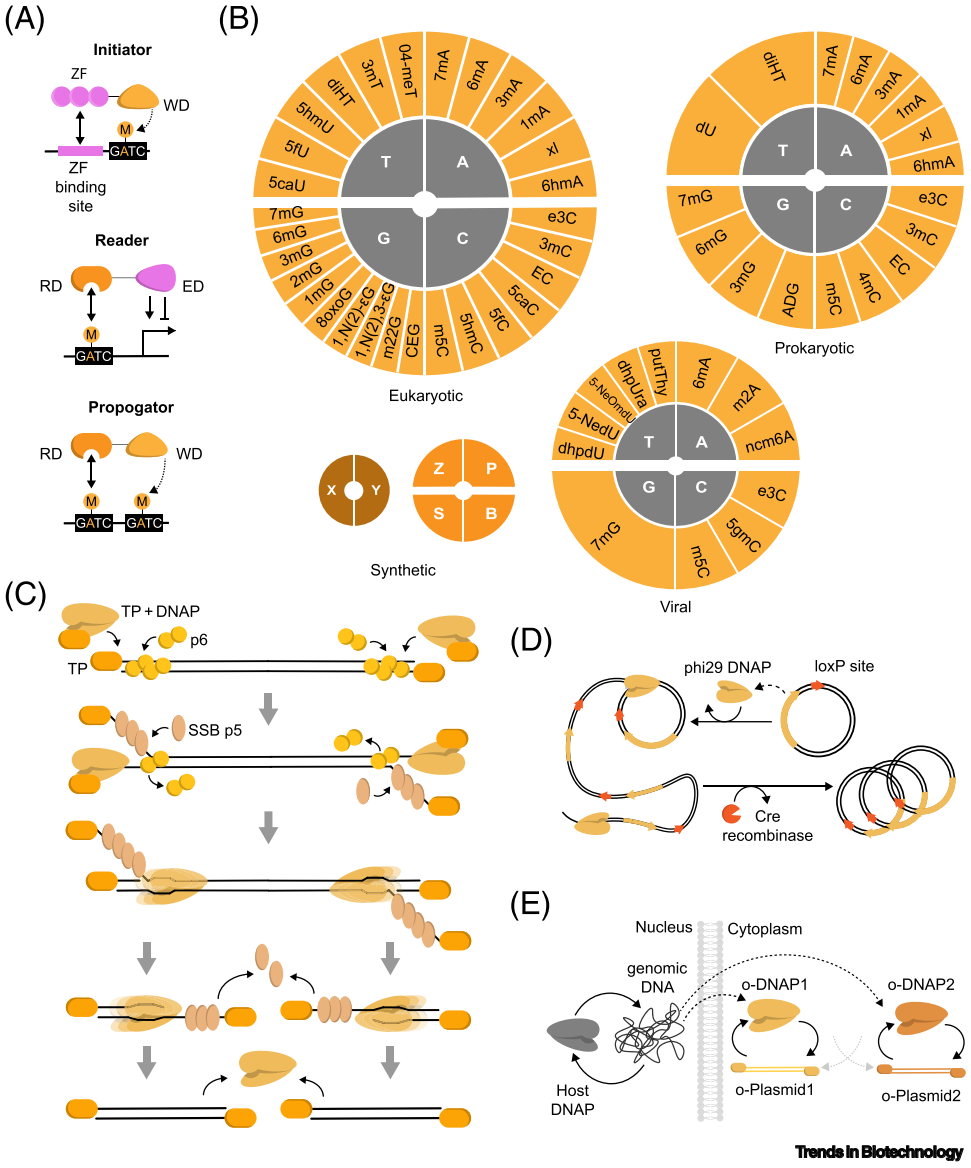

Figure 2. Orthogonal Genetic Information. Schematic representation of components fulfilling aspects of orthogonal genetic information maintenance and propagation. Grayscale elements represent canonical or host-associated components. (A) A system for programable DNA modification composed of two main elements: a writer domain (WD) and reader domain (RD). These can be used to propagate DNA modification and affect gene expression by incorporating an effector domain (ED). (B) Naturally occurring nucleic acid modifications [28], which have been expanded with synthetic nucleobases: X–Y rely on hydrophobic interactions, whereas Z–P and S–B use hydrogen bonding reminiscent of natural Watson–Crick base pairing. (C) An overview of linear ϕ29 bacteriophage genome replication [39]. (D) An orthogonal transcription and translation-coupled DNA replication system (TTcDR) consisting of a circular DNA containing the ϕ29-DNAP gene and a loxP site. ϕ29-DNAP initiates rolling-circle replication on circular DNA, yielding two complementary single-stranded DNAs of tandem repeating sequence. Cre recombinase acts on loxP sites to produce circular products [40]. (E) The yeast-derived orthogonal replication system OrthoRep consists of mutually orthogonal cytosolic DNAP–plasmid pairs. Separation by the nuclear envelope prevents crosstalk with the host genome or DNAPs [44]. Abbreviations: DNAP, DNA polymerase; TP, terminal protein; ZF, zinc finger.

systems. Recent work by Holliger and colleagues extended this observation by creating alkyl phosphonate nucleic acids, showcasing how such modifications can give rise to novel biomolecular interactions in evolved aptamers [38].

Orthogonal DNA Replication

Although non-canonical nucleobases may limit interactions with host machinery, a complementary mechanism leverages orthogonal replication of genetic information. Among well-studied systems, orthogonal replication of the lytic *Bacillus* ϕ29 bacteriophage genome is perhaps the most well understood, and used as a model for protein-primed replication [39]. ϕ29 DNA replication is initiated asynchronously at both ends of the linear genome, wherein a terminal protein (TP)–**DNA polymerase** (DNAP) heterodimer recognizes the protein-capped replication origins (Figure 2C). The DNAP then dissociates from the TP and continues processive elongation coupled to strand displacement. Continuous elongation by two DNAPs gives rise to the complete duplication of the parental strands [39]. Taking inspiration from natural viral replication systems, ϕ29 bacteriophage components have enabled minimally self-replicating DNA systems *in vitro* (Figure 2D). Ichihashi and colleagues developed a minimal transcription- and translation-coupled DNA replication (TTcDR) system wherein a plasmid encoding its own DNAP replicates by rolling circle amplification [40]. Following amplification, Cre recombinase loxP sites facilitate recircularization of the nascent DNA, with rounds of directed evolution used to improve circularization efficiency by 60-fold. Danelon and colleagues [7] recently reported the self-replication of a linear DNA flanked by ϕ29 bacteriophage replication origins in liposome-based synthetic cells. While both systems demonstrate the capacity of ϕ29 bacteriophage components to support an orthogonal DNA replication modality *in vitro*, their extension to cellular systems has not yet been demonstrated, suggesting the potential toxicity of ϕ29 protein production *in vivo*.

Natural auxiliary genetic information has been noted beyond bacteriophage systems, including cytosolic plasmids, mitochondrial, and chloroplast genomes. The latter two genomes exist through symbiotic relationships to their host and have evolved to have share components of their replication machinery. A minimal mitochondrial DNA replication system has been recapitulated *in vitro*, consisting of only of two protein factors: DNA polymerase POLγ and TWINKLE helicase [41]. This system has not yet been extended to a heterologous host, likely a reflection of slow polymerase elongation rate (280 bp/min) and a requirement for two asymmetric origins of replication.

Recently, an orthogonal DNA replication (OrthoRep) system has been established in yeast to realize cellular mutation rates beyond the error catastrophe threshold while mitigating negative consequences on host fitness (Figure 2E) [42–44]. OrthoRep builds on the native cytoplasmic plasmids of *Kluveromyces lactis* to enable orthogonal DNA replication, wherein an orthogonal DNAP is generated from the host genome using native machinery and exclusively replicates the cognate cytoplasmic plasmid. Critically, OrthoRep operates in the cytoplasm, whereas host genome replication occurs in the nucleus, and plasmid propagation relies on protein-primed replication. This system has been elaborated to generate two mutually orthogonal and host-independent DNAPs [43], has strain-independent DNA replication activity [45] and has been used to evolve robust drug resistance and improved enzymatic capabilities [46]. Together, this system highlights the utility of leveraging both non-canonical initiation and compartmentalization to facilitate orthogonal DNA replication.

Orthogonal Transcriptional Regulation

Small Molecule Inducible Transcription Factors

Beyond the development and maintenance of orthogonal DNA, mechanisms for insulating transcriptional and regulatory elements have historically served as the foundation for synthetically controlling gene expression, building upon pioneering studies of inducible systems [47–49]. Rules governing model inducer-responsive transcriptional units have been established and natural regulatory components have been exploited through engineering and directed evolution to

yield synthetic paralogs more suited to their new contexts [5]. Inducible systems classically consist of three components: a protein transcription factor (TF); a DNA motif specifically bound by the TF; and a nontoxic, cell-permeable small molecule that promotes or inhibits the binding of a TF to its cognate DNA motif (Figure 3A). Although not typically identified as orthogonal systems during their development, the success of small molecule-inducible systems has derived from significant efforts to insulate their functionality from host machinery and to limit crosstalk with pre-existing components [5]. This includes the development of nonmetabolizable (e.g., isopropyl β-D-1-thiogalactopyranoside; IPTG) or nontoxic (e.g., anhydrotetracycline; aTc) inducers, elimination of native pathways for inducer detection or catabolism (e.g., lactose and arabinose catabolism), or modulation of TFs to mitigate pleiotropic effects (e.g., IPTG and arabinose crosstalk) [50]. Additionally, to overcome undesirable characteristics that had been previously identified in prokaryotic inducible systems (i.e., noise, low dynamic range, and poor sensitivity), Voigt and colleagues developed a panel of 12 mutually orthogonal TFs using an *in vitro* compartmentalized self-replication (CSR) directed evolution strategy, providing a robust set of orthogonal transcriptional regulators [5].

σ Factors

Natural and engineered TFs often regulate transcriptional outcomes via occlusion of effectors [51]. However, some natural effectors can orthogonally regulate transcriptional output using more direct modalities, and have been used to control gene expression *in vivo*. Bacterial σ factors and anti-σ-factors are essential components of **RNA polymerases** (RNAPs), which differentially affect genomic promoters under variable environmental conditions [52,53]. Natively, the membrane-anchored anti-σ-factors bind their cognate σ factor to inhibit association with the host RNAP holoenzyme. Following an external stimulus, the anti-σ-factor dissociates from its cognate σ factor, thereby allowing for σ–RNAP interaction and initiating transcription from its cognate promoter (Figure 3B) [54]. While many stimuli remain unknown, numerous σ factors show mutual and host orthogonality in *Escherichia coli* [54,55]. Of these, extracytoplasmic function (ECF) σ factors are the smallest and most diverse in terms of sequence [56,57]. Rhodius and colleagues investigated the mutual orthogonality of 86 ECF σ factors from phylogenetically diverse organisms in *E. coli* [54]. Following significant promoter optimization of components derived from GC rich genomes, 20 σ factor–promoter pairs were found to be mutually orthogonal. Notably, ECF σ factor and promoter chimeras expanded this panel to 24 mutually orthogonal pairs. Finally, anti-σ-factors can be used as threshold-gates to minimize leaky expression from orthogonal σ-factor-dependent promoters [54].

Programmable Transcription Factors

Whereas steric occlusion of transcriptional machinery is common regulatory modality in prokaryotic systems, many natural and engineered eukaryotic systems often recruit specific factors to defined DNA motifs to affect transcriptional regulation [51]. Programmable TFs have historically required the expression of orthogonal DNA-binding protein modules to regulate transcriptional circuits, incorporating zinc fingers (ZFs) (Figure 3C) [58] or transcription activator-like effectors (TALEs) (Figure 3D) [59] fused to relevant activating [60–63] or repressive [64] effector domains (EDs). In the past decade, the nuclease-null CRISPR-Cas (dCas) proteins have shown increasingly orthogonal regulation of gene expression, and benefit from utilization of a single protein component alongside engineered guide RNAs (gRNA) (Figure 3E). Programmable dCas-based TFs have served as the basis for multi-input transcriptional enhancers and repressors in mammalian cells [65], owing to the relative ease in programmability as compared to ZFs and TALEs.

Mirroring the utility of orthogonal dCas/sgRNA homologs in eukaryotic gene regulation, Baker and colleagues developed a new approach for orthogonal regulation of gene expression

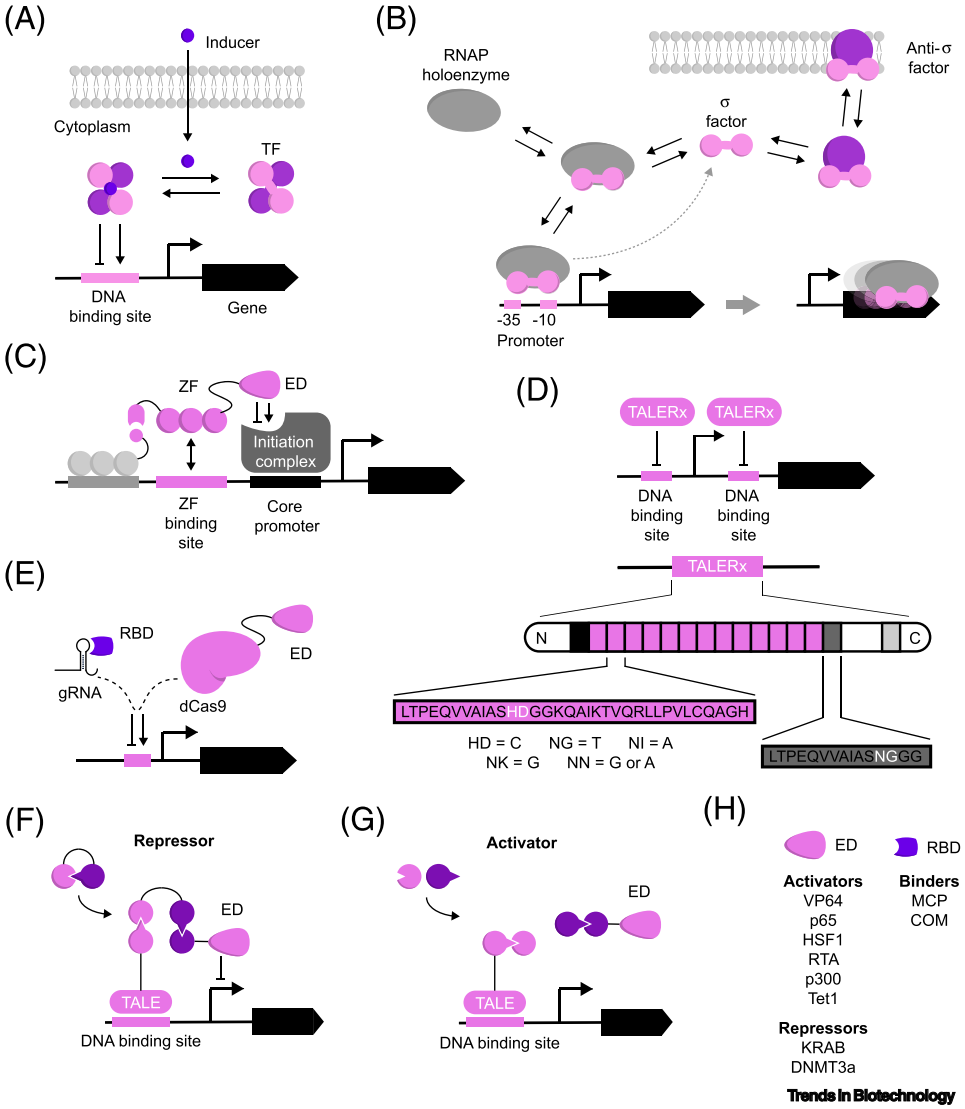

Figure 3. Orthogonal Transcription Components. Schematic representation of components fulfilling aspects of orthogonal transcription. Grayscale elements represent canonical or host-associated components. (A) Small molecule-inducible transcription factors (TF): transcriptional units consist of a TF which binds to its cognate promoter in a small molecule-dependent manner [5]. (B) Extracytoplasmic sigma (σ factors): membrane-anchored anti-σ-factors bind their cognate σ factor to inhibit association with the host RNAP holoenzyme. Following an external stimulus, the σ factor dissociates from its anti-σ-factor, allowing for σ–RNAP interaction and initiating transcription from its cognate promoter [54]. (C) Zinc fingers (ZFs) are modular domains found in many eukaryotic TFs that make sequence-specific contacts with cognate DNA binding sites. Synthetic ZFs have been used as programable TFs to define transcriptional states by fusing it with an effector domain (ED) [58]. (D) TALERx is a programable TF platform based on transcription activator-like effectors (TALEs). Recognition is imparted by two residues in a repeating peptide sequences (pink rectangles with black outline) and correspond to specific nucleotide sequence in a designed DNA sequence [59]. (E) Engineered CRISPR/dCas9-based transcriptional regulators: an engineered guide RNA (gRNA) enables dCas9 binding to the DNA sequence of interest. Either the protein (dCas9) or the gRNA can be modified to include additional functionality by recruiting trans-activating proteins [123]. (F) *De novo* designed logic gates based on heterodimeric protein-protein interaction pairs [66]. These modular heterodimers can be used to generate transcriptional repressors (F) or activators (G). (H) Effector domains and RNA-binding domains (RBDs) used to control gene expression [123].

in vivo through *de novo* protein–protein interaction logic gates [66] (Figure 3F,G). Importantly, these logic gates implement two-component modules that are mutually orthogonal heterodimeric pairs, have tunable binding affinities, and associate cooperatively to limit sensitivity to stoichiometric

input changes. As these heterodimeric pairs interact selectively, this method will likely see facile integration into existing technologies for regulation by small molecules, σ factors, CRISPRs, TALEs, ZFs and/or other effector domains (Figure 3H).

Orthogonal RNAPs and Non-canonical Nucleobases

Naturally derived and engineered prokaryotic and eukaryotic TFs rely on the host transcriptional machinery to regulate gene expression. Compared to these multicomponent systems, bacteriophage-derived RNAPs provide singular polypeptide components that are naturally orthogonal and do not require supplementary cellular factors to affect transcription. **T7** RNAP has proven to be particularly versatile due to its specific recognition of a short 17-bp promoter [67,68], with extensive prokaryotic, eukaryotic, and *in vitro* applications. While additional mutually orthogonal bacteriophage-derived RNAPs are known [69], Ellington and colleagues bolstered this set using compartmentalized partnered replication (CPR) to afford a panel of mutually orthogonal T7 RNAP variants [70]. Additionally, the versatility of bacteriophage RNAPs has been extended to the polymerization of orthogonal nucleobases in live cells [71,72]. Wild-type T7 RNAP has been used to transcribe **mRNAs** and **tRNAs** containing the unnatural nucleotides dNaM (X) and dTPT3 (Y) for an expanded six letter genetic alphabet *in vivo* [71]. The engineered T7-FAL (Y639F, H784A, P266L) variant [73] can support transcription of the eight letter 'hachimoji' code *in vitro* [34], whereas wild-type T7 RNAP can support the polymerization of all but one nucleotide, failing to incorporate riboS opposite deoxyriboB [34]. The extensibility of orthogonal RNAPs to accommodate orthogonal synthetic DNA marks a significant milestone in multilayered orthogonality, and future efforts will likely continue to engineer and/or evolve RNAPs to improve their transcription capabilities for non-canonical nucleobases.

Orthogonal Translation

Post-transcriptional Regulation

Following transcription, mRNAs can be differentially affected via higher order regulatory elements in natural biological systems [74–76], which may influence the functionality of engineered gene circuits. A recent surge in the detection and functional assessment of RNA modifications has revealed roles in dynamic regulation of mRNAs (Figure 4A) [77], suggesting that future engineering efforts may co-opt these modifications to further insulate or affect synthetic circuits. Additionally, regulatory small RNAs (sRNAs) can have profound impacts on mRNA translation (Figure 4B,C). As sRNAs operate via canonical Watson–Crick base pairing, they offer an attractive engineering modality to orthogonally control protein abundance. Indeed, this principle has been recently abstracted to generate synthetic orthogonal RNA regulators of translation in an effort to improve control over protein expression [10,11,78,79]. Recently, Green, Yin, and colleagues developed 15 mutually orthogonal sRNA–mRNA pairs with a dynamic range of up to 300-fold [10]. These devices enable rapid, multi-input Boolean logic gates in *E. coli* that do not require translation of the actuator to function, as is the case in protein-based cellular logic. Similar mechanisms have previously been demonstrated [11,78,79], all of which selectively alter the mRNA 5′ **untranslated region** (UTR) secondary and/or tertiary structure to modulate ribosome binding site (RBS) or start codon availability for translational initiation.

Orthogonal tRNA Engineering

The extensive bioactivities required for tRNA maturation have made their effectors targets for engineering and directed evolution (Figure 4D), and have been aimed largely at improving ncAA incorporation *in vivo* [80]. Broad genetic code expansion efforts have seen the development of orthogonal **aminoacyl-tRNA synthetase** (aaRS)–tRNA pairs to incorporate ncAAs across diverse prokaryotic and eukaryotic cellular systems [81]. Orthogonal aaRS enzymes are well suited to discriminate against native cellular components as each native aaRS–tRNA pair is

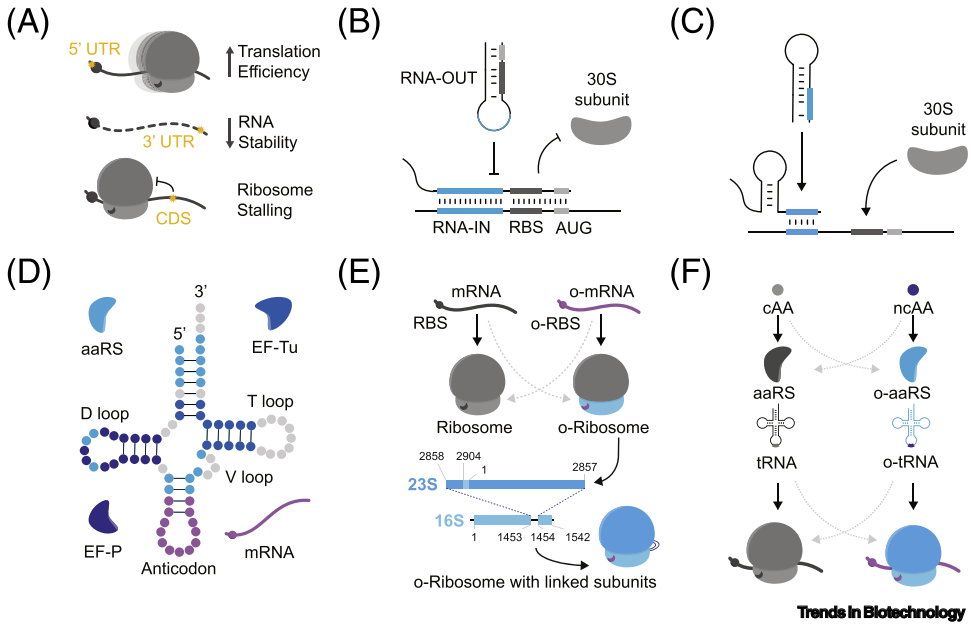

Figure 4. Orthogonal Translation Components. Schematic representation of components fulfilling aspects of orthogonal translation and regulation. Grayscale elements represent canonical or host-associated components. (A) Modified nucleobases found in mRNA can influence translation efficiency and/or mRNA stability [75]. (B) RNA-based translational repressors can be programmed to restrict ribosome initiation *in vivo* [11]. (C) RNA-based translational activators can be engineered to liberate ribosome translation initiation sites *in vivo* [78]. (D) Efforts to engineer tRNAs have affected interactions with various protein-translation factors, further orthogonalizing these biomolecules from host machinery [124]. (E) Orthogonal ribosomes, rely on an o-RBS/o-anti-RBS interaction, can be used to selectively direct ribosome initiation to an orthogonal transcript. Covalent linkage of circularly permutated rRNAs further improves cellular orthogonality [88,89]. (F) Genetic code expansion via non-canonical amino acids (ncAAs): to incorporate ncAAs into proteins, the host is supplemented with an orthogonal aminoacyl-tRNA synthetase (o-aaRS) that selectively charges a cognate orthogonal (o-tRNA) with the ncAA. Abbreviations: EF-P, elongation factor P; EF-Tu, elongation factor Tu; RBS, ribosome binding site; UTR, untranslated region

mutually orthogonal, permitting their directed evolution to catalyze ncAA incorporation. Recently, directed evolution of tRNAs was shown to improve mutual orthogonality for aaRSs in *E. coli* through manipulation of the variable loop region, a known determinant of aaRS–tRNA interactions [82]. Rational engineering of the **elongation factor Tu** (EF-Tu)–tRNA interface can similarly improve ncAA incorporation *in vivo* by improving interactions necessary to shuttle ncAA-charged tRNAs to the ribosome [83]. The **elongation factor P** (EF-P) has been suggested as a candidate for further improvement [84]. These extensive developments have been collated to significantly expand the amino acid diversity *in vivo* by incorporating up to three unique ncAAs using multiple approaches for codon reassignment [85]. Future efforts, which will likely integrate recently demonstrated modes for codon expansion using synthetic bases [86] and mutually orthogonal aaRS–tRNA pairs [87], are expected to enable a larger repertoire of ncAAs to be incorporated into a contiguous polypeptide chain.

Dedicated, mRNA-Specific Orthogonal Ribosomes

The ribosome translates the complete cellular protein repertoire and serves as a signaling nexus to affect cell fate decisions and growth dynamics. Accordingly, inadvertent effects on translation from engineered genetic circuits can have devastating consequences on cellular fitness [14]. To mitigate these effects, a specific and orthogonal ribosome (o-ribosome) population can be generated to exclusively translate researcher-defined orthogonal mRNAs (o-mRNAs) (Figure 4E). **Orthogonal translation** has been developed in prokaryotes and integrates an altered,

orthogonal anti-ribosome binding site (o-antiRBS) in the 16S rRNA to selectively initiate protein translation using a corresponding o-RBS [88]. However, both the orthogonal and native 16S rRNAs can freely associate with wild-type 23S rRNAs, leading to the production of heterogenous ribosome populations *in vivo*. Pioneering work has addressed this limitation by covalently linking circularly permutated rRNAs to yield a contiguous, orthogonal ribosome (Figure 4E) [89,90]. More recently, the association of covalently linked ribosomal subunits with their native counterparts has been minimized by optimizing the intersubunit linker, facilitating the directed evolution of enzymatic capabilities that are inaccessible to otherwise native o-ribosomes [91,92].

The development of increasingly orthogonalized ribosomes *in vivo* may enable expanded functions or alternative mechanisms of translation. In *E. coli*, only three codons are natively used to initiate protein: AUG (83%), GUG (14%), and UUG (3%) [93–95], yet 47 of the 64 native codons may also support translation initiation [96]. Ribosomes engineered or evolved to leverage the remaining codons for translational initiation may therefore impart further degrees of orthogonality *in vivo*. Furthermore, improved modes of orthogonal translation catalyze the incorporation of diverse ncAAs in cellular proteins. There has been significant progress towards this goal over nearly 30 years, with the development of orthogonal aaRSs (o-aaRS) that selectively charge cognate tRNAs with >150 unnatural amino acids (Figure 4F) [97,98]. Historically, these methods co-opt poorly used stop codons (e.g., amber/ochre stop), but have suffered from low efficiencies due to competition with native termination factors [99,100]. To fulfil the potential of ncAA incorporation, complete genome modification [101], *de novo* synthesis of recoded genomes [102], expanded four-base (**quadruplet**) **codons** [6,103–105] and synthetic nucleobases [86] have been successfully used to reassign codons *in vivo*.

Retroactivity and Resource Allocation

Limiting the impact of engineered circuits on the expression and function of endogenous genes or proteins has been a major area of research. Specifically, minimizing the burden of engineered circuits on both the host cell and on one another can have profound impacts on their functionality. Discrete modules, typically characterized in isolation, will often fail to function in concert in novel user-defined contexts. This failure is often attributed to: (i) off-target effects between circuit components and/or with host components [5,106]; (ii) unbalanced flux through a signaling cascade, wherein intermediate products are overproduced above the necessary concentrations [107–110]; and/or (iii) unbalanced host resource allocation to generate the requisite biomolecules, whose production can be coupled due to resource competition (TFs, RNAPs, ribosomes, and metabolites) [13]. Collectively, these factors often manifest as a reduction in host growth rate [19], which can exacerbate engineered circuit function by changing protein dilution [111] or translation rates [112]. Strategies to mitigate these outcomes have incorporated buffering loops that function at more rapid kinetics than the engineered circuits [107], including sRNA-based [12] and phosphorylation-based circuit capacitors [113]. The continued implementation of these or related strategies in higher order engineered circuits will undoubtedly be necessary to support their increasingly complex designs.

Concluding Remarks and Future Perspectives

While significant effort has been devoted to the creation of orthogonal components that alleviate interactions with the host, these innovations have not yet coalesced into a contiguous and orthogonal central dogma (see Outstanding Questions). A major contributing factor has been the development and characterization of orthogonal parts required to achieve a desired outcome, exclusively insulating the necessary elements. For example, the selective control of gene transcription using orthogonal TFs has become routine and typically requires a small number of components [5], whereas the incorporation of numerous ncAAs *in vivo* remains a major challenge

Outstanding Questions

Three orthogonal pairs of nucleobases can be replicated and transcribed *in vitro*, with one pair functioning *in vivo*. Combining all three artificial base pairs would exceed the diversity of possible codon combinations in extant biological nucleic acids. Accordingly, how close are we to genes that are exclusively encoded by orthogonal synthetic nucleotides? Is additional biosynthetic machinery needed to realize this goal?

Given that entire organisms (e.g., *E. coli*) can function with a few hundred transcription factors, how many independent transcriptional regulators are needed to build elaborate and complex synthetic behavior that can mirror natural processes?

Will future synthetic genetic circuit development move towards highly programable components (e.g., CRISPR-based systems), harness the potential orthogonality of heterologous components, rely on *de novo* modular elements, or a combination thereof?

Can polypeptide translation truly be made orthogonal from the host cell given the number of biomolecules involved? Would true orthogonalization require physical separation (e.g., designer membranes, compartmentalization) or can it be achieved more directly?

Orthogonal aaRS/tRNA pairs have enabled ncAA incorporation in living cells, often using recoded organisms that eliminate specific codon-amino acid assignments. How many novel or recoded codons can be used simultaneously in a living cell and what factors limit this number? What enabling technologies are needed to expand the genetic code beyond three ncAAs in a contiguous polypeptide in living cells?

Current synthetic biology efforts still rely heavily on host replication and translation components, yet significant effort has been devoted to limit this dependence. What prompts the first study to bring all levels of orthogonality together within the central dogma?

due to crosstalk with host machinery requiring host engineering [114,115]. Nonetheless, combining orthogonal modalities at multiple nodes of the central dogma has greatly improved circuit performance for otherwise intractable bioactivities. Numerous strategies have improved ncAA incorporation efficiency with orthogonal aaRS–tRNA pairs while limiting host fitness consequences [116]: controlling o-ribosomes with inducible TFs [89,103], developing complete transcription–translation networks that integrate T7 RNAP, o-mRNA, and o-ribosome [117], or even decoding cellular codons bearing synthetic nucleobases [72].

Furthermore, host dependence has restricted the utility of many orthogonal components to a small subset of host organisms. Although some stand-alone orthogonal elements (e.g., T7 RNAP) do not rely on the host machinery and have seen utility in model organisms spanning the tree of life [117–119], most components have been developed to take advantage of host-specific phenomena to direct increasingly orthogonal bioactivities (e.g., RBS/antiRBS complementarity to afford o-ribosomes) [88]. To date, bacterial systems enable the greatest repertoire of orthogonal elements spanning both transcription and translation, while orthogonal DNA replication mechanisms have comparatively lagged. Conversely, robust methods for orthogonal DNA replication and transcription have been described in eukaryotic organisms, whereas orthogonal translation has not seen nearly the same degree of utility as in prokaryotes. These seminal contributions have relied on biological insight or researcher-guided evolution to afford the requisite bioactivities. In the future, it may be possible to develop orthogonal components *de novo*, bolstering the available repertoire of parts and providing access to otherwise intractable cellular functions.

The orthogonalization of engineered gene circuits has reflected a necessity in achieving key bioengineering goals. A fully orthogonal central dogma, therefore, may be required to fulfil increasingly complex user-defined behavior not found in nature. Indeed, the design of advanced cellular circuits to sense, process and actuate in response to a diverse set of input signals will rely not only robust and predictable module behavior, but also the development of guidelines to connect discrete modules for higher order information maintenance and transfer. Progress towards this goal has relied on the accretion of elements and insulation strategies over many decades and from various fields, often limited in their attribution of orthogonal behavior by field-specific jargon or engineering goals. We believe that viewing these components through the lens of orthogonality, whether natural or engineered, most notably as it pertains to insulation at all stages of the central dogma, will catalyze significant progress towards attaining a fully orthogonal, cellular central dogma and enable truly modular and predictable engineered biological systems.

Acknowledgments

This work was supported by the National Institutes of Health Director's Early Independence Award (DP5-OD-024590), the NASA Exobiology Program (NNH17ZDA001NEXO), and the Broad Institute of MIT and Harvard.

References

1. Müller, I.E. *et al.* (2019) Gene networks that compensate for crosstalk with crosstalk. *Nat. Commun.* 10, 1–8
2. Hartwell, L.H. *et al.* (1999) From molecular to modular cell biology. *Nature* 402, C47–C52
3. Cameron, D.E. *et al.* (2014) A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12, 381–390
4. Brophy, J.A.N. and Voigt, C.A. (2014) Principles of genetic circuit design. *Nat. Methods* 11, 508–520
5. Meyer, A.J. *et al.* (2019) *Escherichia coli* "Marionette" strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.* 15, 196–204
6. Anderson, J.C. *et al.* (2004) An expanded genetic code with a functional quadruplet codon. *PNAS* 101, 7566–7571
7. van Nies, P. *et al.* (2018) Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* 9, 1–12
8. Nielsen, A.A.K. *et al.* (2016) Genetic circuit design automation. *Science* 352, aac7341
9. Carlson, P.D. *et al.* (2018) *De novo* design of translational RNA repressors. *bioRxiv* Published online December 19, 2018. https://doi.org/10.1101/501767
10. Kim, J. *et al.* (2019) *De novo*-designed translation-repressing riboregulators for multi-input cellular logic. *Nat. Chem. Biol.* 15, 1173–1182
11. Mutalik, V.K. *et al.* (2012) Rationally designed families of orthogonal RNA regulators of translation. *Nat. Chem. Biol.* 8, 447–454

12. Ceroni, F. *et al.* (2018) Burden-driven feedback control of gene expression. *Nat. Methods* 15, 387
13. Gyorgy, A. *et al.* (2015) Isocost lines describe the cellular economy of genetic circuits. *Biophys. J.* 109, 639–646
14. Qian, Y. *et al.* (2017) Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.* 6, 1263–1272
15. Kurland, C.G. and Dong, H. (1996) Bacterial growth inhibition by overproduction of protein. *Mol. Microbiol.* 21, 1–4
16. Sleight, S.C. and Sauro, H.M. (2013) Visualization of evolutionary stability dynamics and competitive fitness of *Escherichia coli* engineered with randomized multigene circuits. *ACS Synth. Biol.* 2, 519–528
17. Cardinale, S. and Arkin, A.P. (2012) Contextualizing context for synthetic biology – identifying causes of failure of synthetic biological systems. *Biotechnol. J.* 7, 856–866
18. Moser, F. *et al.* (2012) Genetic circuit performance under conditions relevant for industrial bioreactors. *ACS Synth. Biol.* 1, 555–564
19. Borkowski, O. *et al.* (2016) Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.* 33, 123–130
20. Haseltine, E.L. and Arnold, F.H. (2007) Synthetic gene circuits: design with directed evolution. *Annu. Rev. Biophys. Biomol. Struct.* 36, 1–19
21. Ellefson, J.W. *et al.* (2014) Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol.* 32, 97–101
22. Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356
23. Liu, C.C. *et al.* (2018) Toward an orthogonal central dogma. *Nat. Chem. Biol.* 14, 103–106
24. Kittleson, J.T. *et al.* (2011) Rapid optimization of gene dosage in *E. coli* using DIAL strains. *J. Biol. Eng.* 5, 10
25. Gaidukov, L. *et al.* (2018) A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res.* 46, 4072–4086
26. Weigele, P. and Raleigh, E.A. (2016) Biosynthesis and function of modified bases in bacteria and their viruses. *Chem. Rev.* 116, 12655–12687
27. Park, M. *et al.* (2019) Engineering epigenetic regulation using synthetic read-write modules. *Cell* 176, 227–238.e20
28. Sood, A.J. *et al.* (2019) DNAmod: the DNA modification database. *J. Cheminform.* 11, 30
29. Warren, R.A.J. (1980) Modified bases in bacteriophage DNAs. *Annu. Rev. Microbiol.* 34, 137–158
30. Gommers-Ampt, J.H. and Borst, P. (1995) Hypermodified bases in DNA. *FASEB J.* 9, 1034–1042
31. Yang, Z. *et al.* (2011) Amplification, mutation, and sequencing of a six-letter synthetic genetic system. *J. Am. Chem. Soc.* 133, 15105–15112
32. Dhami, K. *et al.* (2014) Systematic exploration of a class of hydrophobic unnatural base pairs yields multiple new candidates for the expansion of the genetic alphabet. *Nucleic Acids Res.* 42, 10235–10244
33. Hirao, I. and Kimoto, M. (2012) Unnatural base pair systems toward the expansion of the genetic alphabet in the central dogma. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 88, 345–367
34. Hoshika, S. *et al.* (2019) Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science* 363, 884–887
35. Feldman, A.W. *et al.* (2018) A tool for the import of natural and unnatural nucleoside triphosphates into bacteria. *J. Am. Chem. Soc.* 140, 1447–1454
36. Pinheiro, V.B. *et al.* (2012) Synthetic genetic polymers capable of heredity and evolution. *Science* 336, 341–344
37. Wang, L. *et al.* (2011) DNA phosphorothioation is widespread and quantized in bacterial genomes. *PNAS* 108, 2963–2968
38. Arangundy-Franklin, S. *et al.* (2019) A synthetic genetic polymer with an uncharged backbone chemistry based on alkyl phosphonate nucleic acids. *Nat. Chem.* 11, 533–542
39. Salas, M. *et al.* (2016) DNA-binding proteins essential for protein-primed bacteriophage ϕ29 DNA replication. *Front. Mol. Biosci.* 3, 37
40. Sakatani, Y. *et al.* (2018) Self-replication of circular DNA by a self-encoded DNA polymerase through rolling-circle replication and recombination. *Sci. Rep.* 8, 1–11
41. Korhonen, J.A. *et al.* (2004) Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J.* 23, 2423–2429
42. Ravikumar, A. *et al.* (2014) An orthogonal DNA replication system in yeast. *Nat. Chem. Biol.* 10, 175–177
43. Arzumanyan, G.A. *et al.* (2018) Mutually orthogonal DNA replication systems in vivo. *ACS Synth. Biol.* 7, 1722–1729
44. Ravikumar, A. *et al.* (2018) Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell* 175, 1946–1957.e13
45. Javanpour, A.A. and Liu, C.C. (2019) Genetic compatibility and extensibility of orthogonal replication. *ACS Synth. Biol.* 8, 1249–1256
46. Zhong, Z. *et al.* (2020) Automated continuous evolution of proteins *in vivo*. *Synthetic Biol.* Published online May 6, 2020. https://doi.org/10.1021/acssynbio.0c00135
47. de Boer, H.A. *et al.* (1983) The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc. Natl. Acad. Sci. U. S. A.* 80, 21–25
48. Skerra, A. (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene* 151, 131–135
49. Guzman, L.M. *et al.* (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* 177, 4121–4130
50. Lee, S.K. *et al.* (2007) Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Appl. Environ. Microbiol.* 73, 5711–5715
51. Hook-Barnard, I.G. and Hinton, D.M. (2007) Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul. Syst. Biol.* 1, 275–293
52. Helmann, J.D. (1999) Anti-sigma factors. *Curr. Opin. Microbiol.* 2, 135–141
53. Campbell, E.A. *et al.* (2008) Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr. Opin. Microbiol.* 11, 121–127
54. Rhodius, V.A. *et al.* (2013) Design of orthogonal genetic switches based on a crosstalk map of σs, anti-σs, and promoters. *Mol. Syst. Biol.* 9, 702
55. Bervoets, I. *et al.* (2018) A sigma factor toolbox for orthogonal gene expression in *Escherichia coli*. *Nucleic Acids Res.* 46, 2133–2144
56. Helmann, J.D. (2002) The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.* 46, 47–110
57. Staroń, A. *et al.* (2009) The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol. Microbiol.* 74, 557–581
58. Khalil, A.S. *et al.* (2012) A synthetic biology framework for programming eukaryotic transcription functions. *Cell* 150, 647–658
59. Li, Y. *et al.* (2015) Modular construction of mammalian gene circuits using TALE transcriptional repressors. *Nat. Chem. Biol.* 11, 207–213
60. Mali, P. *et al.* (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838
61. Chavez, A. *et al.* (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326–328
62. Hilton, I.B. *et al.* (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510–517
63. Xu, X. *et al.* (2016) A CRISPR-based approach for targeted DNA demethylation. *Cell Discov.* 2, 16009
64. Gilbert, L.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451
65. Gao, Y. *et al.* (2016) Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat. Methods* 13, 1043–1049
66. Chen, Z. *et al.* (2020) *De novo* design of protein logic gates. *Science* 368, 78–84
67. Chamberlin, M. *et al.* (1970) New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature* 228, 227–231
68. Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130

69. Temme, K. *et al.* (2012) Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Res.* 40, 8773–8781
70. Meyer, A.J. *et al.* (2015) Directed evolution of a panel of orthogonal T7 RNA polymerase variants for *in vivo* or *in vitro* synthetic circuitry. *ACS Synth. Biol.* 4, 1070–1076
71. Zhang, Y. *et al.* (2017) A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* 551, 644–647
72. Feldman, A.W. *et al.* (2019) Optimization of replication, transcription, and translation in a semi-synthetic organism. *J. Am. Chem. Soc.* 141, 10644–10653
73. Meyer, A.J. *et al.* (2015) Transcription yield of fully 2′-modified RNA can be increased by the addition of thermostabilizing mutations to T7 RNA polymerase mutants. *Nucleic Acids Res.* 43, 7480–7488
74. Tejada-Arranz, A. *et al.* (2020) Bacterial RNA degradosomes: molecular machines under tight control. *Trends Biochem. Sci.* 45, 42–57
75. Zhao, B.S. *et al.* (2017) Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42
76. Filipowicz, W. *et al.* (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9, 102
77. Zaccara, S. *et al.* (2019) Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* 20, 608–624
78. Isaacs, F.J. *et al.* (2004) Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* 22, 841–847
79. Callura, J.M. *et al.* (2012) Genetic switchboard for synthetic biology applications. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5850–5855
80. Hoffman, K.S. *et al.* (2018) Versatility of synthetic tRNAs in genetic code expansion. *Genes (Basel)* 9, 537
81. Melnikov, S. *et al.* (2012) One core, two shells: bacterial and eukaryotic ribosomes. *Nat. Struct. Mol. Biol.* 19, 560–567
82. Willis, J.C.W. and Chin, J.W. (2018) Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nat. Chem.* 10, 831–837
83. Fan, C. *et al.* (2015) Rationally evolving tRNAPyl for efficient incorporation of non-canonical amino acids. *Nucleic Acids Res.* 43, e156
84. Hammerling, M.J. *et al.* (2020) Strategies for *in vitro* engineering of the translation machinery. *Nucleic Acids Res.* 48, 1068–1083
85. Italia, J.S. *et al.* (2019) Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites. *J. Am. Chem. Soc.* 141, 6204–6212
86. Fischer, E.C. *et al.* (2020) New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nat. Chem. Biol.* 16, 570–576
87. Cervettini, D. *et al.* (2020) Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase–tRNA pairs. *Nat. Biotechnol.* Published online April 13, 2020. https://doi.org/10.1038/s41587-020-0479-2
88. Rackham, O. and Chin, J.W. (2005) A network of orthogonal ribosome·mRNA pairs. *Nat. Chem. Biol.* 1, 159–166
89. Orelle, C. *et al.* (2015) Protein synthesis by ribosomes with tethered subunits. *Nature* 524, 119–124
90. Fried, S.D. *et al.* (2015) Ribosome subunit stapling for orthogonal translation in *E. coli*. *Angew. Chem. Int. Ed. Engl.* 54, 12791–12794
91. Schmied, W.H. *et al.* (2018) Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* 564, 444–448
92. Carlson, E.D. *et al.* (2019) Engineered ribosomes with tethered subunits for expanding biological function. *Nat. Commun.* 10, 1–13
93. Adams, J.M. and Capecchi, M.R. (1966) N-formylmethionyl-sRNA as the initiator of protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 55, 147–155
94. Nirenberg, M. and Leder, P. (1964) RNA codewords and protein synthesis. The effect of trinucleotides upon the binding of sRNA to ribosomes. *Science* 145, 1399–1407
95. Blattner, F.R. *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462
96. Hecht, A. *et al.* (2017) Measurements of translation initiation from all 64 codons in E. coli. *Nucleic Acids Res.* 45, 3615–3626
97. Dumas, A. *et al.* (2015) Designing logical codon reassignment – expanding the chemistry in biology. *Chem. Sci.* 6, 50–69
98. Melo Czekster, C. *et al.* (2016) *In Vivo* biosynthesis of a β-amino acid-containing protein. *J. Am. Chem. Soc.* 138, 5194–5197
99. Liu, C.C. and Schultz, P.G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* 79, 413–444
100. Hoesl, M.G. and Budisa, N. (2012) Recent advances in genetic code engineering in *Escherichia coli*. *Curr. Opin. Biotechnol.* 23, 751–757
101. Lajoie, M.J. *et al.* (2013) Genomically recoded organisms expand biological functions. *Science* 342, 357–360
102. Fredens, J. *et al.* (2019) Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 569, 514–518
103. Neumann, H. *et al.* (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441–444
104. Niu, W. *et al.* (2013) An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS Chem. Biol.* 8, 1640–1645
105. Wang, K. *et al.* (2014) Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* 6, 393–403
106. Halleran, A.D. and Murray, R.M. (2018) Cell-free and *in vivo* characterization of Lux, Las, and Rpa quorum activation systems in *E. coli*. *ACS Synth. Biol.* 7, 752–755
107. Grunberg, T.W. and Del Vecchio, D. (2020) Modular analysis and design of biological circuits. *Curr. Opin. Biotechnol.* 63, 41–47
108. Jayanthi, S. *et al.* (2013) Retroactivity controls the temporal dynamics of gene transcription. *ACS Synth. Biol.* 2, 431–441
109. Lee, T.-H. and Maheshri, N. (2012) A regulatory role for repeated decoy transcription factor binding sites in target gene expression. *Mol. Syst. Biol.* 8, 576
110. Del Vecchio, D. *et al.* (2008) Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.* 4, 161
111. Tan, C. *et al.* (2009) Emergent bistability by a growth-modulating positive feedback circuit. *Nat. Chem. Biol.* 5, 842–848
112. Klumpp, S. and Hwa, T. (2014) Bacterial growth: global effects on gene expression, growth feedback and proteome partition. *Curr. Opin. Biotechnol.* 28, 96–102
113. Mishra, D. *et al.* (2014) A load driver device for engineering modularity in biological networks. *Nat. Biotechnol.* 32, 1268–1275
114. Johnson, D.B.F. *et al.* (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.* 7, 779–786
115. Hong, S.H. *et al.* (2014) Cell-free protein synthesis from a release factor 1 deficient Escherichia coli activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol.* 3, 398–409
116. Bryson, D.I. *et al.* (2017) Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat. Chem. Biol.* 13, 1253–1260
117. An, W. and Chin, J.W. (2009) Synthesis of orthogonal transcription-translation networks. *Proc. Natl. Acad. Sci.* 106, 8477–8482
118. Dower, K. and Rosbash, M. (2002) T7 RNA polymerase-directed transcripts are processed in yeast and link 3′ end formation to mRNA nuclear export. *RNA* 8, 686–697
119. Lieber, A. *et al.* (1989) High level gene expression in mammalian cells by a nuclear T7-phase RNA polymerase. *Nucleic Acids Res.* 17, 8485–8493
120. Wei, C.-M. *et al.* (1975) Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. *Cell* 4, 379–386
121. Forchhammer, K. *et al.* (1989) Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature* 342, 453–456
122. Low, S.C. and Berry, M.J. (1996) Knowing when not to stop: selenocysteine incorporation in eukaryotes. *Trends Biochem. Sci.* 21, 203–208
123. Komor, A.C. *et al.* (2017) CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 168, 20–36
124. Mukai, T. *et al.* (2017) Rewriting the genetic code. *Annu. Rev. Microbiol.* 71, 557–577